STANDARD OPERATING PROCEDURE

2341.01A

R7 Geospatial Data Deliverables

August 7, 2015

McLaughlin, Casey

ENST/EDAB

APPROVED:

Lisa Messinger — 8/13/15

Peer Reviewer — Date

Cinda Papio — 9/28/15

ENST Division Director — Date

Beth Miller — 9/30/15

AWMD Division Director — Date

Carl — 9/30/15

CNSL Division Director — Date

Nellie Flournoy — 10/2/15

Acting ARA

PLMG Division Director — Date

Mary P. Peterson — 10/2/2015

SUPR Division Director — Date

Karen A. Flournoy — 10/5/15

WWPD Division Director — Date

Diane Harris — 10/05/2015

Independent Quality Assurance Reviewer — Date

Recertified:

Reviewer					
Date					

Table of Contents

1. Purpose

The purpose of this Standard Operating Procedure is to establish a uniform process for submission and storage of geospatial data to Region 7. This document provides guidance for contractors, grantees, potentially responsible parties, regional staff and others who provide geospatial data to EPA Region 7 programs, projects, or staff.

This document specifies geospatial file delivery formats for all geospatial materials developed in support of geospatial-related work for and within EPA Region 7 to ensure locational data consistency and integrity. It is the intent of EPA Region 7 to acquire, catalog and manage all geospatial files comprehensively across all projects to:

- ensure future use and access to EPA;
- provide an archive of work accomplished;
- apply appropriate data standards and formats;
- maintain and serve data that spatially represent features pertinent to on-going EPA efforts;
- provide a geospatially consistent basis for future activities such as CERCLA Five Year Reviews; and
- attach/incorporate into consent decrees and orders

2. Applicability

This document covers the types of geospatial data deliverables anticipated in Region 7 and how the Region would like to receive these deliverables. Additionally, data standards, formats, and best management practices are identified and should be used. This SOP applies to all personnel, grantees, and potentially responsible parties involved in generation and or/storage of geospatial data for Region 7. *It is the responsibility of those providing deliverables to the Region to adhere to the procedures provided in this document to the best of their abilities.* The Region relies on EPA staff such as grant/contracting officers, On-Scene Coordinators, Remedial Project Managers, and inspectors to ensure data are submitted for long-term storage and use at EPA.

3. Summary of Procedure

Geospatial data should be created using industry best practices, following EPA National Geospatial Data Policy, and be properly documented as specified in national GIS metadata polices and outlined in this SOP.

4. Definitions/Acronyms

CAD stands for Computer Assisted Drawing and is the use of computer systems to assist in the creation, modification, analysis, or optimization of a design.

Esri (formerly known as the Environmental Systems Research Institute or E.S.R.I.) is an international supplier of Geographic Information System (GIS) software, web GIS and geodatabase management applications. Esri is the current EPA enterprise GIS software provider.

FGDC stands for Federal Geographic Data Committee and is an interagency committee that promotes the coordinated development, use, sharing, and dissemination of geospatial data on a national basis.

Geospatial Data means information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth. This information may be derived from, among other things, remote sensing, mapping, and surveying technologies. Statistical data may be included in this definition at the discretion of the collecting agency.

GIS stands for Geographic Information System and is a system designed to capture, store, manipulate, analyze, manage, and present all types of spatial or geographical data.

Machine-Readable refers to information or data that is in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost (https://project-opendata.cio.gov/glossary/).

Metadata is structured information that describes, explains, locates, or otherwise makes it easier to search and retrieve (for example, through a search engine). The simplest definition of metadata is “structured data about data.”

Rasters consist of a matrix of cells (or pixels) organized into rows and columns (or a grid) where each cell contains a value representing information, such as temperature. Rasters are digital aerial photographs, imagery from satellites, digital pictures, or even scanned maps.

Vectors are coordinate-based data models that represents geographic features as points, lines, and polygons. Each point feature is represented as a single coordinate pair, while line and polygon features are represented as ordered lists of vertices. Attributes are associated with each vector feature, as opposed to a raster data model, which associates attributes with grid cells.

5. Personnel Qualifications

Persons should be knowledgeable in GIS mapping and data management best practices and they should be able to demonstrate organizational and communication skills.

6. Procedures

6.1. GIS Formatted Data Files

All final version geospatial-related files acquired or developed to support mapping and/or spatial analysis by a contractor or grantee are considered property of the EPA and are required to be submitted to EPA. This includes, but is not limited to, all GIS, CAD, and image formatted files used to develop maps for any scoping or decision document developed for EPA, as well as any spatial file used to inform a decision on site management or development. Only final versions of each layer are required for delivery to EPA, and must be in an approved format as specified in this SOP. In addition, all electronic geospatial data, whether vector or raster, must be in a well-known projection defined (have a projection defined and embedded in or associated with the data file), and in the case of CAD data must NOT be in page space or a custom site-specific projection; all CAD data must be in known real world coordinate space, ideally conforming to the projection specifications outlined below. Should tabular data be appropriate to connect location information with attribute information, then documentation specifying the primary and foreign keys is required. Please review the EPA National Geospatial Data Policy for additional principals, responsibilities and requirements for collecting and managing geospatial data. Should coordinate information be provided in tabular format, it should contain *at minimum* the following fields (see EPA Latitude/Longitude Data Standard):

- **ID** – a unique identifier given to each feature.
- **Latitude** – the Y coordinate in decimal degrees, 6 decimal places.
- **Longitude** – the X coordinate in decimal degrees (negative), 6 decimal places.
- **Horizontal Datum** – the datum of the coordinates
- **Horizontal Measure –** the horizontal measure, in meters, of the relative accuracy of the latitude and longitude coordinates.
- **Vertical Measure (when collected) –** The measure of elevation (i.e., the altitude), above or below a reference datum.
- **Reference Point Name** – the name that identifies the place (Plant Entrance, Water Well, Storage Tank, Sampling Point, etc…).

Additionally, all static maps that appear in an EPA document should be in an electronic Adobe PDF format with fonts embedded and at a resolution of *300 dots per inch (dpi)* or greater. Finally, any dynamic maps used in final map production, such as Esri ArcMap/Pro documents (.mxd, .aprx), may also be required for delivery to EPA with accompanying data in a stand-alone directory structure. Such documents should to be provided as documented Esri map packages (.mpk). See the EPA Geospatial/Non-Geospatial Metadata Style Guide for best practices on Esri map documents.

6.2. Projection Requirements

All geospatial files submitted to EPA must have spatial reference information that describes the projection/coordinate system, datum, units of measure and where applicable, the collection methods. The EPA requests that all vector data be submitted in **UTM in Zone 14-16 (NAD83) in meters**. Raster data, such as aerial photographs may be submitted in their native projection, and maps should be in the appropriate projection/coordinate system for the area depicted. EPA Region 7 GeoResources Team members will consult and advise on projection, coordinate, and datum details for submission to EPA.

6.3. Metadata Requirements

All geospatial files developed for EPA are required by Executive Order 12906 and EPA National Geospatial Data Policy to have associated metadata. EPA requires FGDC compliant metadata on all geospatial files developed for site support. Region 7 also requires that all dynamic maps (ArcMap documents) have metadata completed. The Content Standard for Digital Geospatial Metadata can be found at www.fgdc.gov. Metadata, including information about the data's projection, can be developed using one of several built-in or add on tools within a GIS, and typically is associated with the geometry file as an XML file. EPA Region 7's Metadata Steward is available to assist with development of required metadata and recommend following the EPA Geospatial/Non-Geospatial Metadata Style Guide (Version1.0) October 7, 2014.

6.4. Organizational Requirement

If the project is complex, a directory structure and readme text file in the upper level directory that describes the structure is required. Because EPA will be managing data across many projects, it is important to make your submittals as understandable as possible. A *recommended* directory structure is as follows:

<**Project_Name**>
- |_ **Docs** (reports, SOPs, correspondence, and other such documents)
- |_ **Maps** (MXDs and PDFs. Map names should use the project name as a prefix)
- |_ **Raster** (aerial photos, satellite imagery, logos, DEMs, and other raster type data)
- |_ **Source** (original unmodified data that may have been acquired from external/internal sources)
- |_ **Tables** (MS-Access databases, spreadsheets, delimited text files, or other such tabular data not stored in a geodatabase)
- |_ **Vectors** (geodatabases, shape files, and other approved vector data formats)

File naming conventions should be logical, consistent, and contain no spaces or special characters. An underscore may be used in lieu of a space. The format for dates is 8 digits in year, month, day order (yyyymmdd).

6.5. Delivery Requirements

EPA will accept data delivered on CD-ROM, DVD, or external hard drive, as well as direct electronic submission via email or FTP site. Other delivery methods may be allowed if those requirements present a significant burden or as technology changes.

Acceptable EPA Geospatial Data Formats:

The following file formats are considered acceptable and all maps and data must include an associated metadata document:

DATA
Vector - -UTM in Zone 14-16 (NAD83) in meters
File Geodatabase (.gdb) *Preferred Shape File (.shp, .shx, .dbf, .prj, .sbx, .sbn) Personal Geodatabase (.mdb) ESRI Map Package (.mpk)
Raster – native projection acceptable
TIFF image with world reference file or as a GeoTIFF (.tif, .tfw) JPEG image with world reference file (.jpg, .jpw) ERDAS Imagine image with pyramid file (.img, .rrd) MrSid image (.sid) ESRI Grid DEM Multi spectral satellite imagery (generally .bil)
TINs – appropriate projection/coordinate system for the area depicted
ESRI TIN
CAD - UTM in Zone 14-16 (NAD83) in meters
DXF layer separates (.dxf) CAD drawing files (.dwg) Microstation and Integraph Design files (.dgn)
Tabular – primary keys should be clearly identified/documented
MS-Access database (.mdb) MS-Excel spreadsheet (.xls, xlsx) Delimited text file (.txt, .csv)

MAPS
Static
Adobe PDF at 300 dpi or better with embedded fonts (.pdf)
Dynamic
ESRI ArcMap/Pro Document (.mxd, .aprx) please specify version ESRI Map Package (.mpk)
FGDC Compliant METADATA (ISO 19115)
XML (.xml)

7. Records Management

The project officer shall be responsible for records management per program-specific policies. Data provided to the GeoResources Team does not replace submittal to the project officer or other records retention schedules. Data submitted to the GeoResources Team is in addition to regular records retention.

8. Quality Assurance/Quality Control

Geospatial data should have documented quality assessments consistent with a data quality objectives and as defined in applicable standards like the EPA Latitude/Longitude Data Standard or a programmatic specific standards. Each dataset or feature as necessary, should include information that describes the accuracy and precision of the collection/creation method. The project officer shall verify that data collected meets the project needs as defined in any data objectives. Geospatial data developed at the Agency shall include horizontal and vertical (when applicable) positional accuracy information for all data sets (see section 4.6, National Geospatial Data Policy Accuracy Tier Table, EPA Geospatial Technical Document Specification).

9. References

EPA Geospatial Metadata Technical Specification Version 1.0, CIO 05-002, November 2, 2007. This document establishes guidelines for publishing geospatial metadata for data sets, applications and services developed by the EPA and applied in EPA's Environmental Dataset Gateway metadata tool.

EPA Geospatial/Non-Geospatial Metadata Style Guide, Version1.0, October 7, 2014. This guide provides suggestions on metadata preparation for EPA geospatial and non-geospatial resources for internal and external sharing. Detailed guidance on metadata requirements, standard language, and naming conventions are provided with the goal of simplifying the documentation process and standardizing (to the extent possible) EPA dataset metadata and documentation.

EPA Geospatial Metadata Technical Specification, Version 1.0, November 2, 2007. This document describes the guidelines for publishing geospatial metadata for data sets, applications, and services.

EPA National Geospatial Data Policy, August, 2008. The National Geospatial Data Policy (NGDP) establishes principles, responsibilities, and requirements for collecting and managing geospatial data used by Federal environmental programs and projects within the jurisdiction of the U.S. EPA.

Global Positioning System Technical Implementation Guidance, September 2003. This document serves as reference guidance for Agency staff using GPS equipment and for individuals responsible for the maintenance of this equipment.

Institutional Control Vector Profile Technical Specification, EX000015.1, January 6, 2006. This specification establishes the requirements for documenting geographic lines and boundary coordinates and related method, accuracy, and description data for places of interest to the EPA. Latitude/Longitude Data Standard, EX000017.2, January 6, 2006. This data standard establishes the requirements for documenting latitude and longitude coordinates and related method, accuracy, and description data for places of interest to the EPA.

National Geospatial Data Policy Procedure for Geospatial Metadata Management, CIO 2131-P-01-0, October 2010. The purpose of this document is to establish procedures, requirements and responsibilities for maintaining geospatial metadata used by EPA programs.